

Teadmiste formaliseerimine

aka

Knowledge Representation
(and Reasoning)

KR & R

Intro

Course page

http://lambda.ee/wiki/Teadmiste_formaliseerimine

Course theme

SQL & conventional software <----> natural language

- Structured facts: SQL
- Unstructured knowledge: natural language
- *Focus of the course: semi-structured knowledge*

Course theme

SQL & conventional software <----> natural language

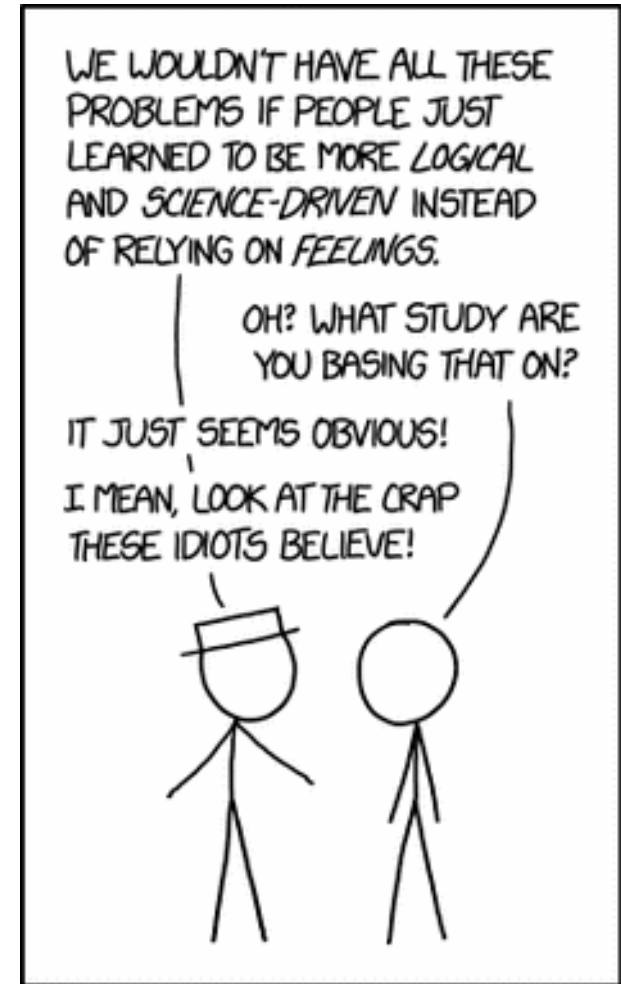
- Structured facts: SQL
- Unstructured knowledge: natural language
- *Focus of the course: semi-structured knowledge*

Capturing meaning in natural language sentences:

understanding what is said, answering questions, explaining answers, calculating plans of action.

Question answering and inference using NL

- Strongly related to commonsense reasoning.
- Long history, little success with symbolic methods. CYC!?
- Large Language Model (LLM) explosion: [success with issues](#)
 - Explainability
 - Hallucinations
 - Hard to connect with tools like databases, computation, etc
 - Hard to do deep inference
 - Controllability: how to add a lot of rules?
 - etc



Tiny GPT examples

- Finds the correct answer :

If an animal likes honey, then it is probably a bear. Most bears are big, although young bears are not big. John is an animal who likes honey. Mike is a young bear. Who is big?

- Does not find the correct answer:

If an animal likes honey, then it is probably a bear. Most bears are big, although young bears are not big. John is an animal who likes honey. Mike is a young bear. Mike can eat a lot. Penguins are birds who cannot fly. John took the block from the colored table. The table was really nice. The robot arm lifted a blue block from the table. Who is big?

Mis on Tallinnast Helsingisse üle vee jalutamise kiirusrekord?

ChatGPT:

Tallinnast Helsingisse üle vee jalutamise kiirusrekordi kohta mul teavet ei ole, kuna see ei ole tavaline ega praktiline viis kahe linna vahel liikumiseks.

Tallinna ja Helsingi vahemaa Soome lahe kohal on umbes 80 kilomeetrit, mis teeks selle väga keeruliseks, kui mitte võimatuks, kauguseks jalgsi liikuda.

Tavaliselt reisivad inimesed Tallinna ja Helsingi vahel laevadega, kiirlaevadega või lennukitega, mis pakuvad kiiremaid ja tõhusamaid ühendusi. Soome lahe ületamine jalutades oleks äärmiselt ohtlik ja ebapraktiline ettevõtmine, mida ei soovitata ega tohiks teha.

Mille poolest on tuntud Konstantin Päts?

ChatGPT:

Konstantin Päts (1874–1956) oli Eesti poliitik ja riigitegelane, kes mängis olulist rolli Eesti iseseisvuse taastamisel ja riigi varajases ajaloos. Ta on tuntud mitmel viisil:
Eesti iseseisvuse taastaja: Konstantin Päts oli üks peamisi Eesti iseseisvuse taastamise liidreid pärast Esimese maailmasõja lõppu ja Vabadussõda (1918-1920). Ta mängis olulist rolli Eesti iseseisvuse väljakuulutamisel 1918. aastal ja Eesti Vabariigi esimese valitsuse moodustamisel.
Eesti Vabariigi president: Pärast iseseisvuse taastamist valiti Konstantin Päts Eesti Vabariigi esimeseks presidendiks 1938. aastal. **Ta oli ka president aastatel 1938-1940 ning siis veel kord pärast Eesti iseseisvuse taastamist 1992. aastal.**

Poliitiline juht ja reformaator: Konstantin Päts oli oluline poliitiline juht Eesti Vabariigi varases ajaloos ja aitas kaasa mitmetele sotsiaalsetele ja majanduslikele reformidele. Ta juhtis Eestit ka rasketel aegadel, sealhulgas okupeerimise ajal Nõukogude Liidu poolt ja hiljem **Saksa okupatsiooni ajal Teise maailmasõja käigus.**

Can we fix end-to-end learning?

Nobody knows so far.

A lot of people put their hope on hybrid systems:

Combine learning with rule-based reasoning, arithmetic, SQL etc.

Journey: main steps of the course

SQL and logic

SQL and triples / RDF / JSON-LD

Simple rules like in RDFa

Large semi-structured knowledge bases

Machine learning in NLP: vector representation of words

Machine learning in NLP: large language models (BERT, GPT)

Reasoning with uncertain knowledge / uncertain rules

Semantic parsing

Hybrid question answering: machine learning + rule-based reasoning

Wikipedia

In artificial intelligence (AI), [commonsense reasoning](#) is a human-like ability to make presumptions about the type and essence of ordinary situations humans encounter every day.

These assumptions include judgments about the nature of physical objects, taxonomic properties, and peoples' intentions.

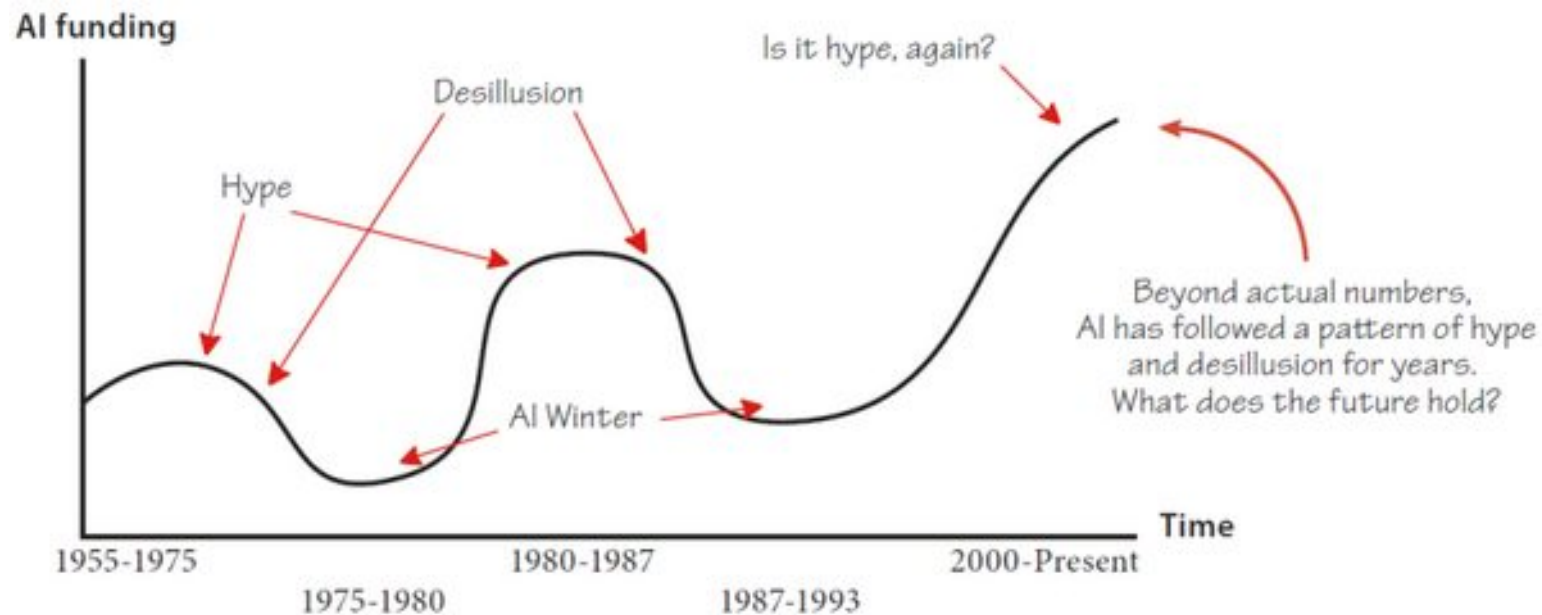
Wikipedia AGI

Artificial general intelligence (AGI) is the ability of an intelligent agent to understand or learn any intellectual task that a human being can. It is a primary goal of some artificial intelligence research and a common topic in science fiction and futures studies.

AGI can also be referred to as **strong AI**, **full AI**, or **general intelligent action**, although some academic sources reserve the term "strong AI" for computer programs that experience sentience or consciousness.

What is pure classical KR & R?

- A classic old-school A.I. subfield *for actual human-like thinking*
- Used to set lofty goals extremely hard to achieve (and did not)
- Famous A.I. winters:



Approach to thinking in KR&R

Symbolic A.I:

- Write down (human?) knowledge using rules and facts.
- Answer questions and solve problems by deriving new knowledge using rules.

Next, small examples using an automated reasoner for first order logic:
run these yourself in <http://logictools.org>

spacity(kuressaare).

railroad(tallinn,tapa).

railroad(tapa,tartu).

highway(tallinn,virtsu).

sealane(virtsu,kuivastu).

highway(kuressaare,kuivastu).

railroad(X,Y) => railroad(Y,X).

highway(X,Y) => highway(Y,X).

sealane(X,Y) => sealane(Y,X).

railroad(X,Y) => easytravel(X,Y,use(train,X,Y)).

highway(X,Y) => easytravel(X,Y,use(bus,X,Y)).

sealane(X,Y) => easytravel(X,Y,use(ship,X,Y)).

easytravel(X,Y,P1) & easytravel(Y,Z,P2) => easytravel(X,Z,combine(P1,P2)).

easytravel(tartu,X,Y) & spacity(X) => \$ans(X,Y).

- Replace *spacity(kuressaare)* with

city(kuressaare).

has(kuressaare,coast).

hasnrof(kuressaare,spa)=4.

population(kuressaare)=10000.

\$less(population(X),20000) => smallcity(X).

(city(X) & has(X,coast) & \$less(3,hasnrof(X,spa)) & smallcity(X)) => spacity(X).

...

Note: it can be pretty hard to find solutions

- Finding a solution is semidecidable: no guarantee that we can show that a solution does *not* exist.
- Even small problems can be hard. A solvable tiny example from group theory, ca 57 steps in a proof (try it out on logictools.org):

$\text{multiply}(A, \text{inverse}(\text{multiply}(B, \text{multiply}(C, \text{multiply}(\text{multiply}(\text{inverse}(C), \text{inverse}(\text{multiply}(D, B))), A)))) = D.$

$\text{multiply}(\text{inverse}(a1), a1) \neq \text{multiply}(\text{inverse}(b1), b1).$

A narrow subfield of KR &R: automated reasoning

- AKA “Automated Theorem Proving”
- Focus on developing algorithms and systems for automatically solving problems written in logic.
- Yearly competitions <http://www.tptp.org/CASC/> and the main conference CADE <http://www.cadeinc.org/conferences>
- Focus more on mathematical kinds of problems, not so much common sense or very large knowledge bases.
- In other words, top systems work best with relatively few rules and deep complex proofs.
- Not too many high-performance systems (like, my system Gandalf was world leading at the end of 20th century for many years, and the current GKC system is one of the top systems nowadays)

What happens with this approach?

- Small toy problems are easy to present and solve.
- Presenting and solving large complex problems for common-sense reasoning has turned out to be extremely hard.

Why is symbolic reasoning with NL hard?

- **Semantic parsing**, i.e. translating natural language to logic, is extremely hard due to the highly complex and exception-rich nature of natural language.
- **Existing knowledge bases of “common sense”** do not cover a critical mass of the basic understanding of the world even a small child possesses.
- **Classical first order reasoning** itself cannot cope with contradictory knowledge items, probabilistic or uncertain information and exceptions to rules.
- **Finding logic-based proofs** often requires long derivations and the huge knowledge base causes a quick combinatorial explosion of the search space.

Hypothesis: ML and LLMs can alleviate all the main issues above

Famous case study: CYC



Wikipedia

- **Cyc** is a long-term artificial intelligence project that aims to assemble a comprehensive ontology and knowledge base that spans the basic concepts and rules about how the world works.
- Douglas Lenat began the project in July 1984 ... since January 1995, has been under active development by the **Cycorp** company.
- The Cyc project has been described as "[one of the most controversial endeavors of the artificial intelligence history](#)". Catherine Havasi, CEO of Luminoso, says that Cyc is the predecessor project to IBM's Watson. Machine-learning scientist Pedro Domingos refers to the project as a "catastrophic failure".

A mostly failed limited approach: Semantic Web

- Idea: make it easy to publish, harvest, integrate and use structured knowledge on the web.
- Proposals worked out: RDF data model and OWL rule-like language
- RDF main idea: represent all data in a single three-column table of <object id> <property> <value> triplets like this:

client_1 name „John Brown“

client_1 balance 200

What happened with the Semantic Web?

Negative:

- Mostly academic research leading to over-complicating even simple things.
- People were not really keen to publish data or rules on the web.

Positive:

- Attempt to make it a bit simpler in a closely related area: linked data
- JSON-LD language giving RDF interpretation to JSON
- <http://schema.org> repository of suggested type/property names

Positive case studies of KR & R

- Google's Knowledge Graph used e.g. for infoboxes and Google Assistant
- Facebook Knowledge Graph
- Apple Siri knowledge base
- IBM Watson question-answering system from 2011

You may want to read a big overview of knowledge graphs:

<https://arxiv.org/pdf/2003.02320.pdf>

IBM Watson winning Jeopardy! in 2011



Example questions from Jeopardy with Watson

- It's just a bloody nose! You don't have this hereditary disorder once endemic to European royalty : *Haemophilia*
- You just need a nap! You don't have this sleep disorder that can make sufferers nod off while standing up : *Narcolepsy*
- Heitor Villa-Lobos dedicated his "12 Etudes" for this instrument to Andres Segovia : *Guitar*
- Paganini's "24 Capricci" set the standard for etudes for this instrument : *Violin*
- Rembrandt's Biblical Scene "Storm on the Sea of" this was stolen from a Boston museum in 1990 : *Galilee*

Neurosymbolic / hybrid systems

- A new area, much smaller than end-to-end learning
- Most of the research investigates ways to use rules while learning
- Our direction: start with the symbolic reasoner and find ways to add learning and integrations with LLM systems
- What we have: a symbolic pipeline for commonsense reasoning with NL, geared towards experimenting with ML and LLM augmentation

Most landmark A.I. systems are hybrid

... except for object recognition and story/picture generation

- Siri and Google Assistant
- Most automated driving systems
- AlphaGo and AlphaZero
- AlphaFold

Can we pre-build some necessary structure?

Like making the system internally know things like:

- There are objects, and they are mostly persistent.
- Objects are in space and time and follow some basic physics.
- Objects have properties.
- Objects have relations between each other.
- Relations have some basic logical properties.
- Some things are dangerous, some are desirable

...

Hybrid systems: ML + rules

Two approaches

- Inject rules into a machine learning model (extremely active field)
- Add machine learning to a rule-based system (only a few groups)

Inject rules into machine learning

- Immense number of possible ways to do that
- People are actively experimenting
- Typically we get small gains over pure end-to-end transformers

Example: GPT simplifying Wikipedia

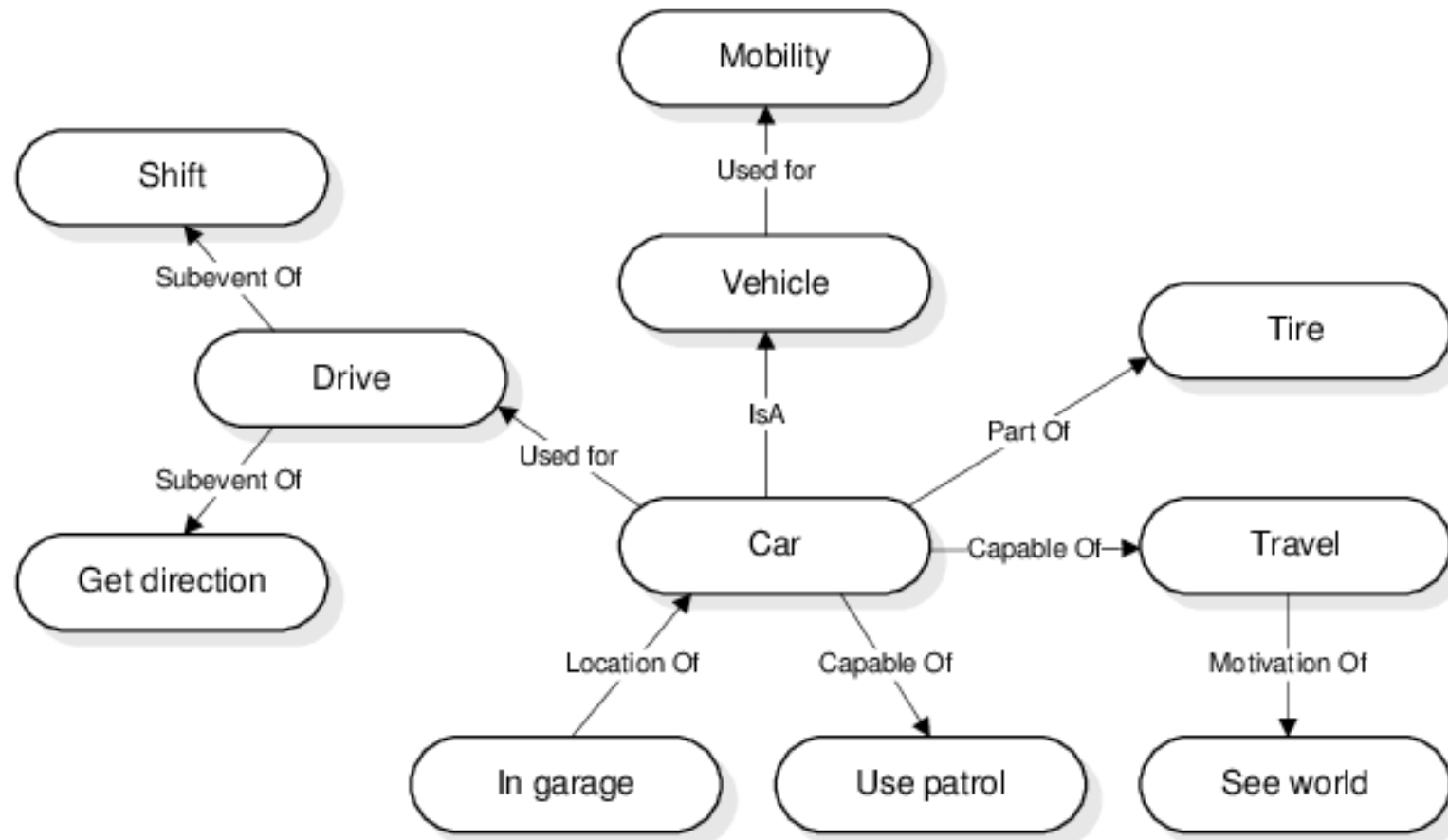
- **Prompt:** Simplify, maximally shorten and split the sentence after colon to shortest possible separate subsentences, to make it understandable for children. Replace pronouns like 'they', 'it', 'he', 'she' in the result with nouns and proper nouns present in the result, like 'Birds can fly. Birds have feathers.' instead of 'Birds can fly. They have feathers.'
- **Wikipedia sentence given after prompt:** Penguins are a group of aquatic flightless birds from the order Sphenisciformes of the family Spheniscidae.
- **GPT gives sentences we can parse:** Penguins are special birds. They live in water. Penguins cannot fly. They are from the Sphenisciformes order. They belong to the Spheniscidae family.

One big issue: good commonsense KBs?

Some of the most well-known:

- wordnet
- Dbpedia
- wikidata
- yago
- babelnet
- conceptnet
- atomic
- nell
- framenet
- cyc
- sumo
- TPTP
- schema.org

Typical KG contents: simple triplets like



Examples of research in hybrid systems

<https://arxiv.org/abs/2301.13379>

<https://deepmind.google/discover/blog/alphageometry-an-olympiad-level-ai-system-for-geometry/>